Linear Optimal Estimation in Feedback Hybrid Systems with Application to Tracking in Clutter

Daniel Sigalov, Tomer Michaeli, and
Yaakov Oshman, *Fellow, IEEE*

***Abstract*—A generalized state space representation of a dynamical system with random modes is presented. The new formulation includes a term, in the dynamics equation, which depends on the most recent state's linear minimum mean squared error (LMMSE) estimate. This can be used to model the behavior of a feedback control system featuring a state estimator. The measurement equation is allowed to depend on the previous LMMSE estimate of the state, which can be used to represent the fact that measurements are obtained from a validation window centered at the predicted measurement and not from the entire surveillance region. The matrices comprising the system's mode constitute an independent stochastic process. The proposed formulation generalizes several problems considered in the past, and allows a unified modeling of new ones. The LMMSE optimal filter is derived for the considered problem. The approach is demonstrated in the context of target tracking in clutter and is shown to yield performance that is competitive to that of several popular nonlinear methods.**

***Index Terms*—State estimation, target tracking, fault detection and isolation, clutter and data association, hybrid systems.**

I. Introduction

State estimation in dynamical systems with randomly switching coefficients is an important problem in a variety of applications. The most natural examples are maneuvering target tracking and fault detection and isolation (FDI) algorithms, featured, e.g., in aerospace navigation systems. The standard modeling presumes that the dynamics of the continuously-valued state, and, possibly, its measurement equation, are controlled by an evolving mode that takes discrete values. This is the well known concept of hybrid systems [1].

Various problems have been formulated using a hybrid state space framework. In problems involving uncertain, or intermittent observations, such as [2]–[5], the mode affects the matrices of the measurement equation. In maneuvering target tracking applications, considered in, e.g., [6]–[9], the mode usually affects the matrices of the dynamics equation.

We consider a state space representation of dynamical systems with randomly switching coefficients, that is more general than the commonly used one in two major aspects. First, we allow the system input to depend on the latest estimate of the state, as is common practice in closed loop control systems. In this work, the state estimate is taken to be the linear minimum mean squared error (LMMSE) estimate. Likewise, we allow the measurement equation to depend on the latest state estimate (taken as well to be the LMMSE estimate). This can be used to represent the fact that observations are not taken in the entire admissible space, but rather in a small validation window set about the predicted measurement of the state.

It is well known [7] that, even for the simplest case of independently switching modes, the optimal estimate of the state cannot be obtained without resorting to exhaustive enumeration. Therefore, significant efforts have been dedicated to developing suboptimal approaches for state estimation in hybrid systems and especially for the important subclass of jump linear systems (JLS). Originally derived in [7], the generalized pseudo-Bayesian (GPB) filter approximates the conditional expectation of the state of a Markov JLS (MJLS) by pruning the measurement history. Perhaps the most successful suboptimal nonlinear method is the interacting multiple model (IMM) algorithm, proposed in [8], in which history pruning is accompanied by a wise cooperation of the primitive Kalman filters comprising the algorithm, thus improving the performance over GPB with only a slight increase in computational compexity. Alternatively, one may consider optimality within a narrower family of linear filters. While [2], [3] and others considered special cases of problems involving random modes, De Koning [10] derived a Kalman filter (KF) like algorithm for the most general case of JLS with independent random modes. Costa [11] derived a linear optimal scheme for the most general case of a standard MJLS by means of state augmentation. Recently, it was shown in [12] that, in some cases, parts of the state may be estimated optimally while others in a linear optimal manner.

In this paper we consider optimality within the family of linear filters for a generalized state space representation of dynamical systems. One of the attractive properties of LMMSE estimators is their minimax optimality [13]. That is, they attain the lowest worst-case MSE, in comparison to any other (even nonlinear) method, for the same first- and second-order moments of the underlying distributions. We derive a linear optimal filter that may be conveniently implemented in a recursive form, eliminating the need for unbounded memory, and without state augmentation. Our filter reduces to previously reported results when the parameters of the underlying problem are appropriately adjusted.

As an illustration of the approach, we show how the problem of tracking a target in clutter may be formulated within our framework. We show that in this setting, the performance of our filter is competitive to that of several classical nonlinear methods.

The remainder of this paper is organized as follows. In Section II we define and discuss the proposed state space modeling approach and survey some related contributions. The linear optimal recursive state estimation algorithm is developed in Section III. An application of the approach to target tracking in clutter, followed by a representative numerical study, is presented in Section IV. Concluding remarks are provided in Section V.

II. System Model and Related Work

We consider the dynamical system

$$x_{k+1} = A_k x_k + B_k u_k + C_k w_k \tag{1a}$$

$$y_k = H_k x_k + G_k v_k + F_k \hat{x}_{k-1}, \tag{1b}$$

in which $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^m$ are the state and measurement vectors at time k, respectively. The noise processes $\{w_k\}$ and $\{v_k\}$ constitute zero-mean unity-covariance white sequences, and x_0 is a random vector with mean $\bar{x}_0$ and second-order moment P_0.

We consider two versions for the modeling of the input vector u_k. In the first (standard) case, u_k is a known deterministic input. However, as in some cases u_k serves as a closed loop control signal, it is common practice to let it depend on the most recent estimate of the state. Thus, in the second variant we set $u_k = \hat{x}_k$, where $\hat{x}_k$ is the LMMSE estimate of x_k using the measurement history $\mathcal{Y}_k \triangleq \{y_1, \dots, y_k\}$.

Likewise, the term $\hat{x}_{k-1}$ in the measurement equation is the LMMSE estimate of x_{k-1} based on the measurement history $\mathcal{Y}_{k-1}$. Affecting the measurement at time k, the term $F_k \hat{x}_{k-1}$ can be used to represent the fact that observations are not taken in the entire (feasible) space, but, rather, in a small (admissible) validation window, set about the predicted measurement.

Finally, the system mode, $\mathcal{M}_k \triangleq \{A_k, B_k, C_k, H_k, G_k, F_k\}$, constitutes an independent random process whose distribution at time k, $\mathrm{p}(\mathcal{M}_k)$, is known. The random quantities $\{w_k\}$, $\{v_k\}$, $\{\mathcal{M}_k\}$, and x_0 are assumed to be mutually independent.

We seek to obtain the linear optimal estimate $\hat{x}_{k+1}$ using the measurements $\mathcal{Y}_{k+1}$. It will be shown in the sequel that $\hat{x}_{k+1}$ in

our setting conveniently possesses the recursive form

$$\hat{x}_{k+1} = L_k \hat{x}_k + K_k y_{k+1} + J_k u_k. \tag{2}$$

Namely, the desired linear optimal solution will be shown to fuse the (extrapolated) previous estimate with the most recently acquired measurement vector, effectively eliminating the need to store the entire measurement sequence. For the case in which $u_k = \hat{x}_k$, the terms $L_k \hat{x}_k$ and $J_k \hat{x}_k$ in (2) may be grouped together.

Note that the described problem does not require the system mode to assume values in a discrete domain. In addition, the above formulation allows evolution not only of the entries of the matrices constituting the mode, but also of their dimensions. This observation provides a convenient basis for treating problems that, to the best of the our knowledge, have not been previously considered, as will be discussed in detail in Section IV.

For the setting in which no feedback terms are present (namely, $B_k = 0$ and $F_k = 0$), several variants and special cases of the presented problem have been considered in the past. Nahi [2] devised a linear optimal estimator having the recursive form (2) for the problem of state estimation in the presence of multiplicative measurement faults. In this setting, only the matrix H_k in (2) is random. A generalization of Nahi's work has been recently proposed in [5], where multiplicative faults were accompanied by additive ones representing measurement biases. This can be cast within the formulation (2), where only the matrices H_k and G_k are random. De Koning [10] considered the most general case of independently switching modes and Costa [11] developed, by means of state augmentation, a recursive implementation of the LMMSE filter for systems with modes obeying Markov dynamics taking values in some discrete domain. Several additional contributions related to the considered problem include [3], in which a generalized version of Nahi's problem was addressed by allowing correlated fault indicators; [4], that allowed correlations between subsequent fault variables; [6], that proposed a linear optimal estimator for the static multiple model (SMM) problem [14]; and [15], that considered linear optimal estimation in systems with bounded and unbounded numbers of actuator faults. Nonlinear suboptimal solutions for related problems were proposed in [7]–[9], [16] and references therein.

III. Linear Optimal Recursive Estimation

We now derive the LMMSE recursive filter for the case where u_k is a known deterministic signal. The result for the stochastic case is given in Section III-E.

Let Y_k be the random vector obtained by concatenating the measurements comprising $\mathcal{Y}_k$. The desired recursive form (2) will be derived using the following lemma, which follows directly from [17, p. 190] and from the linearity of the MMSE estimator in the Gaussian case.

Lemma 1. *Let x, y and z be random vectors and let $\hat{x}(z)$ and $\hat{x}(y,z)$ denote, respectively, the LMMSE estimates of x using z, and using both y and z. In addition, let $\hat{y}(z)$ denote the LMMSE estimate of y using z. Then*

$$\hat{x}(y,z) = \hat{x}(z) + \Gamma_{x\tilde{y}} \Gamma_{\tilde{y}\tilde{y}}^{-1} \tilde{y}, \tag{3}$$

where $\tilde{y} = y - \hat{y}(z)$ and Γ_{xy} is the cross-covariance matrix between the random vectors x and y.

Letting $z \triangleq Y_k$ and $y \triangleq y_{k+1}$ and using Lemma 1, we obtain the following form for the LMMSE estimate of $\hat{x}_{k+1}$ using $\mathcal{Y}_{k+1}$:

$$\hat{x}_{k+1} = \hat{x}_{k+1}^- + \Gamma_{x_{k+1}\tilde{y}_{k+1}} \Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1} \tilde{y}_{k+1}, \tag{4}$$

where $\hat{x}_{k+1}^-$ is the LMMSE estimate of x_{k+1} using the measurements $\mathcal{Y}_k = \{y_1, \ldots, y_k\}$, $\tilde{y}_{k+1} \triangleq y_{k+1} - \hat{y}_{k+1}^-$, and $\hat{y}_{k+1}^-$ is the LMMSE estimate of y_{k+1} using $\mathcal{Y}_k$. It follows that

$$\hat{x}_{k+1}^- = \mathbb{E}\left[A_k\right] \hat{x}_k + \mathbb{E}\left[B_k\right] u_k \tag{5}$$

$$\begin{aligned}\hat{y}_{k+1}^- &= \mathbb{E}\left[H_{k+1}\right] \hat{x}_{k+1}^- + \mathbb{E}\left[F_{k+1}\right] \hat{x}_k \\ &= (\mathbb{E}\left[H_{k+1}\right] \mathbb{E}\left[A_k\right] + \mathbb{E}\left[F_{k+1}\right]) \hat{x}_k + \mathbb{E}\left[H_{k+1}\right] \mathbb{E}\left[B_k\right] u_k.\end{aligned} \tag{6}$$

Plugging (5) in (4) we identify the desired matrix coefficients K_k, L_k, and J_k of (2) as follows:

$$K_k = \Gamma_{x_{k+1}\tilde{y}_{k+1}} \Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1} \tag{7}$$

$$L_k = (I - K_k \mathbb{E}\left[H_{k+1}\right]) \mathbb{E}\left[A_k\right] - K_k \mathbb{E}\left[F_{k+1}\right] \tag{8}$$

$$J_k = (I - K_k \mathbb{E}\left[H_{k+1}\right]) \mathbb{E}\left[B_k\right]. \tag{9}$$

We now compute the covariance terms $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$ and $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$.

A. Computation of $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$

Since $\hat{y}_{k+1}^-$ is unbiased,

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}\left[(x_{k+1} - \mathbb{E}\left[x_{k+1}\right])(y_{k+1} - \hat{y}_{k+1}^-)^\top\right] \\ &= \mathbb{E}\left[x_{k+1}(y_{k+1} - \hat{y}_{k+1}^-)^\top\right],\end{aligned} \tag{10}$$

Substituting (1b) and (6),

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} = \mathbb{E}&\left[x_{k+1}(H_{k+1}x_{k+1} + G_{k+1}v_{k+1} + F_{k+1}\hat{x}_k)^\top\right] \\ &- \mathbb{E}\left[x_{k+1}((\mathbb{E}\left[H_{k+1}\right] \mathbb{E}\left[A_k\right] + \mathbb{E}\left[F_{k+1}\right])\hat{x}_k)^\top\right] \\ &- \mathbb{E}\left[x_{k+1}(\mathbb{E}\left[H_{k+1}\right] \mathbb{E}\left[B_k\right] u_k)^\top\right].\end{aligned} \tag{11}$$

Utilizing the independence of x_{k+1} and v_{k+1}, and canceling out identical terms (11) becomes

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} = \mathbb{E}[x_{k+1}x_{k+1}^\top]\mathbb{E}[H_{k+1}^\top] - \mathbb{E}[x_{k+1}\hat{x}_k^\top]\mathbb{E}[A_k^\top]\mathbb{E}[H_{k+1}^\top] \\ - \mathbb{E}[x_{k+1}]u_k^\top \mathbb{E}[B_k^\top]\mathbb{E}[H_{k+1}^\top].\end{aligned} \tag{12}$$

Before proceeding, we define the following matrices

$$\Sigma_k \triangleq \mathbb{E}[x_k x_k^\top] \tag{13}$$

$$\Lambda_k \triangleq \mathbb{E}[\hat{x}_k \hat{x}_k^\top] = \mathbb{E}[\hat{x}_k x_k^\top] \tag{14}$$

$$\Upsilon_k \triangleq \mathbb{E}[x_k] u_k^\top = \mathbb{E}[\hat{x}_k] u_k^\top \tag{15}$$

$$\Delta_k \triangleq u_k u_k^\top, \tag{16}$$

where the RHS of (14) follows from the orthogonality principle satisfied by $\tilde{x}_k$, the estimation error of x_k, and the RHS of (15) from the unbiasedness of $\hat{x}_k$. Note that Σ_k, Λ_k, and Δ_k are symmetric.

Now, utilizing the independence of $\hat{x}_k$ and w_k,

$$\begin{aligned}\mathbb{E}\left[x_{k+1}\hat{x}_k^\top\right] &= \mathbb{E}\left[(A_k x_k + B_k u_k + C_k w_k)\hat{x}_k^\top\right] \\ &= \mathbb{E}\left[(A_k x_k + B_k u_k)\hat{x}_k^\top\right] \\ &= \mathbb{E}\left[A_k\right] \Lambda_k + \mathbb{E}\left[B_k\right] \Upsilon_k^\top,\end{aligned} \tag{17}$$

which yields for (12)

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} = \Sigma_{k+1}&\mathbb{E}[H_{k+1}^\top] \\ &- (\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top)\mathbb{E}[A_k^\top]\mathbb{E}[H_{k+1}^\top] \\ &- \mathbb{E}[x_{k+1}]u_k^\top \mathbb{E}[B_k^\top]\mathbb{E}[H_{k+1}^\top].\end{aligned} \tag{18}$$

To express $\mathbb{E}\left[x_{k+1}\right]$ in terms of $\mathbb{E}\left[x_k\right]$ we substitute (1a)

$$\begin{aligned}\mathbb{E}\left[x_{k+1}\right] &= \mathbb{E}\left[A_k x_k + B_k u_k + C_k w_k\right] \\ &= \mathbb{E}\left[A_k\right] \mathbb{E}\left[x_k\right] + \mathbb{E}\left[B_k\right] u_k.\end{aligned} \tag{19}$$

Using the latter in (18) and rearranging the summands, the final expression for $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$ reads

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} &= \Sigma_{k+1}\mathbb{E}[H_{k+1}^\top] - \big(\mathbb{E}[A_k](\Lambda_k\mathbb{E}[A_k^\top] + \Upsilon_k\mathbb{E}[B_k^\top]) \\ &\quad + \mathbb{E}[B_k](\Upsilon_k^\top\mathbb{E}[A_k^\top] + \Delta_k\mathbb{E}[B_k^\top])\big)\mathbb{E}[H_{k+1}^\top]. \qquad (20)\end{aligned}$$

B. Computation of $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$

Recall that $\hat{y}_{k+1}^-$ is the LMMSE estimate of y_{k+1} and, consequently, $\tilde{y}_{k+1}$ is orthogonal to $\hat{y}_{k+1}^-$. Hence,

$$\begin{aligned}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}[(y_{k+1} - \hat{y}_{k+1}^-)y_{k+1}^\top] \\ &= \mathbb{E}[y_{k+1}y_{k+1}^\top] - \mathbb{E}[\hat{y}_{k+1}^- y_{k+1}^\top] \\ &= \mathbb{E}[y_{k+1}y_{k+1}^\top] - (\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}])\mathbb{E}[\hat{x}_k y_{k+1}^\top] \\ &\quad - \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k\mathbb{E}[y_{k+1}^\top], \qquad (21)\end{aligned}$$

where we have utilized the independence of $\{\hat{x}_k, y_{k+1}\}$ and $\{A_k, B_k, H_{k+1}, F_{k+1}\}$. Using (1b) and the independence of $\{\hat{x}_k, x_{k+1}\}$, $\{H_{k+1}, G_{k+1}, F_{k+1}\}$ and v_{k+1}, we have that

$$\begin{aligned}\mathbb{E}[\hat{x}_k y_{k+1}^\top] &= \mathbb{E}\left[\hat{x}_k(H_{k+1}x_{k+1} + G_{k+1}v_{k+1} + F_{k+1}\hat{x}_k)^\top\right] \\ &= \mathbb{E}\left[\hat{x}_k(H_{k+1}x_{k+1} + F_{k+1}\hat{x}_k)^\top\right] \\ &= \mathbb{E}[\hat{x}_k x_{k+1}^\top]\mathbb{E}[H_{k+1}^\top] + \Lambda_k\mathbb{E}[F_{k+1}^\top], \qquad (22)\end{aligned}$$

which, using (17), becomes

$$\begin{aligned}\mathbb{E}[\hat{x}_k y_{k+1}^\top] &= \Lambda_k(\mathbb{E}[A_k^\top]\mathbb{E}[H_{k+1}^\top] + \mathbb{E}[F_{k+1}^\top]) \\ &\quad + \Upsilon_k\mathbb{E}[B_k^\top]\mathbb{E}[H_{k+1}^\top]. \qquad (23)\end{aligned}$$

Due to the independence of $\{x_{k+1}, \hat{x}_k\}$, v_{k+1}, and $\{H_{k+1}, G_{k+1}\}$

$$\begin{aligned}&\mathbb{E}[y_{k+1}y_{k+1}^\top] \\ &= \mathbb{E}[(H_{k+1}x_{k+1} + G_{k+1}v_{k+1} + F_{k+1}\hat{x}_k) \\ &\qquad \times (H_{k+1}x_{k+1} + G_{k+1}v_{k+1} + F_{k+1}\hat{x}_k)^\top] \\ &= \mathbb{E}[H_{k+1}x_{k+1}x_{k+1}^\top H_{k+1}^\top] + \mathbb{E}[G_{k+1}v_{k+1}v_{k+1}^\top G_{k+1}^\top] \\ &\quad + \mathbb{E}[F_{k+1}\hat{x}_k\hat{x}_k^\top F_{k+1}^\top] + \mathbb{E}[H_{k+1}x_{k+1}\hat{x}_k^\top F_{k+1}^\top] \\ &\quad + \mathbb{E}[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top]. \qquad (24)\end{aligned}$$

Consider, for example, the term $\mathbb{E}\left[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top\right]$. From the smoothing property of the conditional expectation,

$$\begin{aligned}\mathbb{E}[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top] &= \mathbb{E}\left[\mathbb{E}[F_{k+1}\hat{x}_k x_{k+1}^\top H_{k+1}^\top \mid F_{k+1}, H_{k+1}]\right] \\ &= \mathbb{E}\left[F_{k+1}\mathbb{E}[\hat{x}_k x_{k+1}^\top \mid F_{k+1}, H_{k+1}]H_{k+1}^\top\right] \\ &= \mathbb{E}\left[F_{k+1}\mathbb{E}[\hat{x}_k x_{k+1}^\top]H_{k+1}^\top\right], \qquad (25)\end{aligned}$$

where the conditioning on H_{k+1} and F_{k+1} was omitted due to the independence of $\{x_{k+1}, \hat{x}_k\}$ and $\{H_{k+1}, F_{k+1}\}$.

Similarly, bearing in mind that $\mathbb{E}\left[x_{k+1}x_{k+1}^\top\right] = \Sigma_{k+1}$, $\mathbb{E}\left[v_{k+1}v_{k+1}^\top\right] = I$, and $\mathbb{E}\left[\hat{x}_k\hat{x}_k^\top\right] = \Lambda_k$, we obtain:

$$\mathbb{E}[H_{k+1}x_{k+1}x_{k+1}^\top H_{k+1}^\top] = \mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^\top] \qquad (26)$$
$$\mathbb{E}[G_{k+1}v_{k+1}v_{k+1}^\top G_{k+1}^\top] = \mathbb{E}[G_{k+1}G_{k+1}^\top] \qquad (27)$$
$$\mathbb{E}[F_{k+1}\hat{x}_k\hat{x}_k^\top F_{k+1}^\top] = \mathbb{E}[F_{k+1}\Lambda_k F_{k+1}^\top]. \qquad (28)$$

For future reference, we also note that

$$\mathbb{E}[A_k x_k x_k^\top A_k^\top] = \mathbb{E}[A_k\Sigma_k A_k^\top] \qquad (29)$$
$$\mathbb{E}[A_k x_k u_k^\top B_k^\top] = \mathbb{E}[A_k\Upsilon_k B_k^\top] \qquad (30)$$
$$\mathbb{E}[B_k u_k u_k^\top B_k^\top] = \mathbb{E}[B_k\Delta_k B_k^\top] \qquad (31)$$
$$\mathbb{E}[C_k w_k w_k^\top C_k^\top] = \mathbb{E}[C_k C_k^\top]. \qquad (32)$$

Substituting (17) in (25), and using (25)-(28) in (24)

$$\begin{aligned}\mathbb{E}[y_{k+1}y_{k+1}^\top] &= \mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^\top] + \mathbb{E}[G_{k+1}G_{k+1}^\top] \\ &\quad + \mathbb{E}[F_{k+1}\Lambda_k F_{k+1}^\top] \\ &\quad + \mathbb{E}\left[H_{k+1}(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top)F_{k+1}^\top\right] \\ &\quad + \mathbb{E}\left[F_{k+1}(\Lambda_k\mathbb{E}[A_k^\top] + \Upsilon_k\mathbb{E}[B_k^\top])H_{k+1}^\top\right]. \qquad (33)\end{aligned}$$

In addition, we obtain in a straightforward manner

$$\begin{aligned}\mathbb{E}[y_{k+1}] &= (\mathbb{E}[H_{k+1}]\mathbb{E}[A_k] + \mathbb{E}[F_{k+1}])\mathbb{E}[x_k] \\ &\quad + \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]u_k. \qquad (34)\end{aligned}$$

Using (26), (27), and (28) in (33), and substituting (23), (33), and (34) in (21) we obtain the final expression for $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$.

$$\begin{aligned}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}} &= \mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^\top] + \mathbb{E}[G_{k+1}G_{k+1}^\top] \\ &+ \mathbb{E}[F_{k+1}\Lambda_k F_{k+1}^\top] - \mathbb{E}[F_{k+1}]\Lambda_k\mathbb{E}[F_{k+1}^\top] \\ &- \mathbb{E}[H_{k+1}]\mathbb{E}[A_k](\Lambda_k\mathbb{E}[A_k^\top] + \Upsilon_k\mathbb{E}[B_k^\top])\mathbb{E}[H_{k+1}^\top] \\ &- \mathbb{E}[H_{k+1}]\mathbb{E}[B_k](\Upsilon_k^\top\mathbb{E}[A_k^\top] + \Delta_k\mathbb{E}[B_k^\top])\mathbb{E}[H_{k+1}^\top] \\ &+ \mathbb{E}\left[H_{k+1}(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top)F_{k+1}^\top\right] \\ &+ \mathbb{E}\left[F_{k+1}(\Lambda_k\mathbb{E}[A_k^\top] + \Upsilon_k\mathbb{E}[B_k^\top])H_{k+1}^\top\right] \\ &- \mathbb{E}[H_{k+1}]\mathbb{E}[A_k]\Lambda_k\mathbb{E}[F_{k+1}^\top] - \mathbb{E}[F_{k+1}]\Lambda_k\mathbb{E}[A_k^\top]\mathbb{E}[H_{k+1}^\top] \\ &- \mathbb{E}[F_{k+1}]\Upsilon_k\mathbb{E}[B_k^\top]\mathbb{E}[H_{k+1}^\top] - \mathbb{E}[H_{k+1}]\mathbb{E}[B_k]\Upsilon_k^\top\mathbb{E}[F_{k+1}^\top]. \qquad (35)\end{aligned}$$

C. Computation of the Second-Order Moments

The second-order moment matrix of x_{k+1} is given by

$$\begin{aligned}\Sigma_{k+1} &= \mathbb{E}[x_{k+1}x_{k+1}^\top] \\ &= \mathbb{E}\left[(A_k x_k + B_k u_k + C_k w_k)(A_k x_k + B_k u_k + C_k w_k)^\top\right] \\ &= \mathbb{E}[A_k\Sigma_k A_k^\top] + \mathbb{E}[A_k\Upsilon_k B_k^\top] + \mathbb{E}[B_k\Upsilon_k^\top A_k^\top] \\ &\quad + \mathbb{E}[B_k\Delta_k B_k^\top] + \mathbb{E}[C_k C_k^\top], \qquad (36)\end{aligned}$$

where we have utilized the independence of x_k, w_k and $\{A_k, B_k, C_k\}$, and (29)–(32).

Next consider Λ_{k+1}. Using (2) we have:

$$\begin{aligned}\Lambda_{k+1} &= \mathbb{E}\left[(L_k\hat{x}_k + K_k y_{k+1} + J_k u_k)x_{k+1}^\top\right] \\ &= (L_k + K_k\mathbb{E}[F_{k+1}])\mathbb{E}[\hat{x}_k x_{k+1}^\top] \\ &\quad + K_k\mathbb{E}[H_{k+1}]\Sigma_{k+1} + J_k u_k\mathbb{E}[x_{k+1}^\top].\end{aligned}$$

Using (17) the latter becomes

$$\begin{aligned}\Lambda_{k+1} &= L_k(\Lambda_k\mathbb{E}[A_k^\top] + \Upsilon_k\mathbb{E}[B_k^\top]) + J_k(\Upsilon_k^\top\mathbb{E}[A_k^\top] + \Delta_k\mathbb{E}[B_k^\top]) \\ &\quad + K_k(\mathbb{E}F_{k+1}(\Lambda_k\mathbb{E}A_k^\top + \Upsilon_k\mathbb{E}B_k^\top) + \mathbb{E}[H_{k+1}]\Sigma_{k+1}). \qquad (37)\end{aligned}$$

Finally,

$$\Upsilon_{k+1} = \mathbb{E}[x_{k+1}]u_{k+1}^\top. \qquad (38)$$

Note that Δ_k is known and does not need to be computed recursively.

D. Algorithm Summary

The complete algorithm comprises the following steps:

Initialization: $\hat{x}_0 = \bar{x}_0$, $\Sigma_0 = P_0 + \bar{x}_0\bar{x}_0^\top$, $\Lambda_0 = \bar{x}_0\bar{x}_0^\top$, $\Upsilon_0 = \bar{x}_0 u_0^\top$, $\Delta_0 = u_0 u_0^\top$.

Algorithm 1

Input: y_{k+1}, u_{k+1}, $\hat{x}_k$, $\mathbb{E}[x_k]$, Σ_k, Λ_k, Υ_k, Δ_k

1: Compute $\mathbb{E}[A_k]$, $\mathbb{E}[B_k]$, $\mathbb{E}\left[C_k C_k^\top\right]$, $\mathbb{E}\left[A_k \Sigma_k A_k^\top\right]$, $\mathbb{E}\left[A_k \Upsilon_k B_k^\top\right]$, and $\mathbb{E}\left[B_k \Delta_k B_k^\top\right]$.
2: Compute $\mathbb{E}[x_{k+1}]$ and Σ_{k+1} using Eqs. (19) and (36).
3: Compute $\mathbb{E}[H_{k+1}]$, $\mathbb{E}\left[G_{k+1} G_{k+1}^\top\right]$, $\mathbb{E}[F_{k+1}]$, $\mathbb{E}\left[F_{k+1} \Lambda_k F_{k+1}^\top\right]$, $\mathbb{E}\left[H_{k+1} \Sigma_{k+1} H_{k+1}^\top\right]$, and $\mathbb{E}\left[H_{k+1}(\mathbb{E}[A_k]\,\Lambda_k + \mathbb{E}[B_k]\,\Upsilon_k^\top) F_{k+1}^\top\right]$.
4: Compute $\Gamma_{x_{k+1}\tilde{y}_{k+1}}$ and $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ using Eqs. (20) and (35).
5: Compute K_k, L_k, and J_k using Eqs. (7), (8), and (9).
6: Compute Λ_{k+1} and Υ_{k+1} using Eq. (37) and (38).
7: Compute $\hat{x}_{k+1}$ using Eq. (2).

Output: $\hat{x}_{k+1}$, $\mathbb{E}[x_{k+1}]$, Σ_{k+1}, Λ_{k+1}, Υ_{k+1}

Recursion: For $k = 1, 2, \ldots$ perform the routine of Algorithm 1. Notice that since the distribution of the mode at each time is known, the expectations of steps 1 and 3 of the algorithm may be calculated in a straightforward manner. For example, if the mode variables take discrete values, these expectations reduce to probability-weighted summations of the corresponding matrices. However, in some cases, as demonstrated in Section IV, closed form expressions exist for the above expectations.

E. Random Inputs

In the second variant of (1a), in which $u_k = \hat{x}_k$, it turns out that the roles played by A_k and B_k are identical. Specifically, after replacing u_k with $\hat{x}_k$, at each step of the derivation of Section III, A_k and B_k are multiplied by the same quantities. Thus, the filter for the modified problem is obtained from the one described in Alg. 1 by replacing A_k with $A_k + B_k$ and nullifying u_k and Υ_k. An alternative derivation, based on the orthogonality principle, may be found in [18].

IV. Application to Target Tracking in Clutter

In this section we demonstrate the proposed concept by casting the classical problem of tracking in clutter within our formulation, and applying the LMMSE filter of Section III.

A. System and Clutter Models

Consider a single target obeying a linear dynamical model. The evolution of the state is obtained from (1a) by setting $A_k = A$, $B_k = 0$, and $C_k = C$, resulting in

$$x_{k+1} = A x_k + C w_k. \tag{39}$$

Here A and C are deterministic matrices, accounting for the state dynamics and process noise covariance, respectively.

At time k, the target state is observed via the the linear equation

$$y_k^{\mathrm{true}} = H_{\mathrm{nom}} x_k + G_{\mathrm{nom}} v_k^{\mathrm{true}}, \tag{40}$$

where v_k^{true} represents measurement noise. In addition to the measurement y_k^{true}, at each time, a number of clutter detections are obtained. These will be denoted as $\{y_{k,i}^{\mathrm{cl}}\}_{i=1}^{N-1}$, where N is total number of detections. Clutter measurements originate from false (or ghost) targets and do not carry any information about the target of interest. They are, however, indistinguishable from true detections. At each time, the clutter measurements are assumed to be independent of each other, of the clutter measurements at other times, and of the true state and observation. In addition, we assume that these measurements are uniformly distributed in space. As an illustration, we present in Fig. 1 a single realization of a one dimensional scenario in which the true target position is accompanied by the actual measurements and clutter observations.

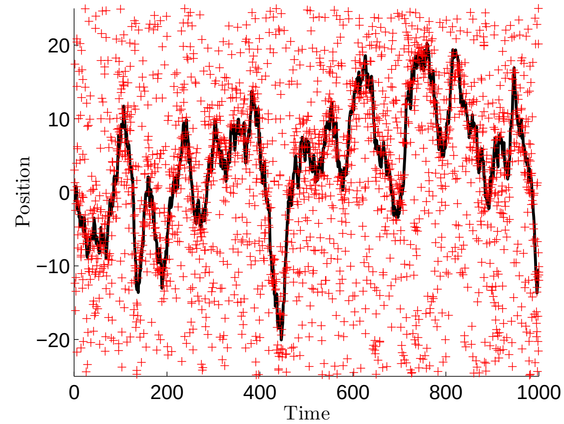

Fig. 1: True target position (solid line) and all the obtained measurements (pluses) vs. time.

To correctly model the distribution of the clutter detections, we note that typically, at each scan, the sensor initiates a validation window centered at the predicted target position, and the algorithm processes only those measurements obtained within the window. Since the clutter detections are uniformly distributed in space, they are also uniformly distributed within the validation window.

We define the measurement vector y_k at time k to be the concatenation of all measurements, $N-1$ of which correspond to clutter, and one originating from the true target. The location of the true measurement within this concatenated vector is, of course, unknown to the algorithm. This setting can be modeled using (1b) by letting the mode $\mathcal{M}_k$ be distributed as

$$\mathcal{M}_k = \{H_k, G_k, F_k\} \tag{41}$$
$$= \begin{cases} \left\{ \begin{pmatrix} H_{\mathrm{nom}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathrm{diag}\begin{pmatrix} G_{\mathrm{nom}} \\ G_{\mathrm{cl}} \\ \vdots \\ G_{\mathrm{cl}} \end{pmatrix}, \begin{pmatrix} 0 \\ H_{\mathrm{nom}} A \\ \vdots \\ H_{\mathrm{nom}} A \end{pmatrix} \right\}, & \text{w.p. } \frac{1}{N} \\ \quad\vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots & \\ \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ H_{\mathrm{nom}} \end{pmatrix}, \mathrm{diag}\begin{pmatrix} G_{\mathrm{cl}} \\ \vdots \\ G_{\mathrm{cl}} \\ G_{\mathrm{nom}} \end{pmatrix}, \begin{pmatrix} H_{\mathrm{nom}} A \\ \vdots \\ H_{\mathrm{nom}} A \\ 0 \end{pmatrix} \right\}, & \text{w.p. } \frac{1}{N}, \end{cases}$$

where $\mathrm{diag}\begin{pmatrix} A_1 & A_2 & \cdots & A_N \end{pmatrix}$ stands for a block diagonal matrix. Here, G_{cl} is the square-root of the covariance matrix associated with the clutter, which, as mentioned above, is uniformly distributed in the validation window.

The first realization of $\{H_k, G_k, F_k\}$ in (41), for example, corresponds to the scenario in which the first of the N observations is the true target measurement, y_k^{true}, generated according to (40), while the other $N-1$ measurements are clutter, each of which is generated according to

$$y_{k,i}^{\mathrm{cl}} = H_{\mathrm{nom}} A \hat{x}_{k-1} + G_{\mathrm{cl}} v_{k,i}^{\mathrm{cl}}, \quad i = 2, \ldots, N. \tag{42}$$

Here, $H_{\mathrm{nom}} A \hat{x}_{k-1}$ is the predicted true measurement at time k, which is also the center of the validation window set by the sensor, so that clutter measurements acquired at time k are uniformly distributed

around this quantity. The overall number of measurements in the validation window, N, is assumed to be known, but may vary in time. Thus, the dimensions of H_k, G_k, and F_k may depend on k.

Notice that we assumed, for simplicity, that the true measurement is always present in the validation window. To account for the possibility that the true measurement does not fall in the validation window, the option

$$\{H_k, G_k, F_k\} = \{\mathbf{0}, I_N \otimes G_{\text{cl}}, \mathbf{1}_N \otimes H_{\text{nom}}A\}, \tag{43}$$

needs to be added to set of possible realizations of $\{H_k, G_k, F_k\}$ appearing in (41). Here, the symbol $\otimes$ stands for the Kronecker product, $\mathbf{1}_N$ is an $N \times 1$ vector comprising all ones, and I_N is the $N \times N$ identity matrix.

Finally, notice that if $N = 0$, namely, there are no measurements in the validation window, then (2) becomes, at the absence of u_k, $\hat{x}_{k+1} = L_k\hat{x}_k$, which corresponds to a simple prediction (time update) without consecutive measurement update, as could be expected.

B. Matrix Computations

To invoke the algorithm presented in Section III we need to compute the application-dependent terms outlined in Steps 1 and 3 of Alg. 1. Although these may be evaluated numerically, via direct summations, in the present example closed-form expressions exist, as we show next.

Since the matrices of the dynamics equation are deterministic, $\mathbb{E}[A_k] = A$, $\mathbb{E}[B_k] = 0$, $\mathbb{E}\left[C_kC_k^\top\right] = CC^\top$, $\mathbb{E}\left[A_k\Upsilon_kB_k^\top\right] = 0$, $\mathbb{E}\left[B_k\Delta_kB_k^\top\right] = 0$, and $\mathbb{E}\left[A_k\Sigma_kA_k^\top\right] = A\Sigma_kA^\top$. Further, according to the probability distribution defined in (41),

$$\mathbb{E}[H_{k+1}] = \frac{1}{N}\mathbf{1}_N \otimes H_{\text{nom}} \tag{44a}$$

$$\mathbb{E}[F_{k+1}] = \frac{N-1}{N}\mathbf{1}_N \otimes H_{\text{nom}}A. \tag{44b}$$

Computation of the remaining expectations yields

$$\mathbb{E}[H_{k+1}\Sigma_{k+1}H_{k+1}^\top] = \frac{1}{N}I_N \otimes H_{\text{nom}}\Sigma_{k+1}H_{\text{nom}}^\top, \tag{45}$$

$$\mathbb{E}[G_{k+1}G_{k+1}^\top] = \frac{1}{N}I_N \otimes \Big(G_{\text{nom}}G_{\text{nom}}^\top + (N-1)G_{\text{cl}}G_{\text{cl}}^\top\Big), \tag{46}$$

and

$$\mathbb{E}\left[F_{k+1}\Lambda_kF_{k+1}^\top\right] = \Xi \otimes \Big(H_{\text{nom}}A\Lambda_kA^\top H_{\text{nom}}^\top\Big), \tag{47}$$

where Ξ is defined by

$$\Xi = \begin{cases} 0, & N = 1 \\ \frac{1}{N}\left((N-2)\mathbf{1}_N\mathbf{1}_N^\top + I_N\right), & N > 1. \end{cases} \tag{48}$$

Finally,

$$\begin{aligned}\mathbb{E}\left[H_{k+1}(\mathbb{E}[A_k]\Lambda_k + \mathbb{E}[B_k]\Upsilon_k^\top)F_{k+1}^\top\right] &= \mathbb{E}\left[H_{k+1}A\Lambda_kF_{k+1}^\top\right] \\ &= \frac{1}{N}(\mathbf{1}_N\mathbf{1}_N^\top - I_N) \otimes \Big(H_{\text{nom}}A\Lambda_kA^\top H_{\text{nom}}^\top\Big).\end{aligned} \tag{49}$$

Following our assumption, the spatial distribution of the clutter measurements is uniform in the validation window. Thus, $G_{\text{cl}}G_{\text{cl}}^\top$ is the covariance matrix of a random vector uniformly distributed in the validation window. For computational simplicity, we assume that the window is rectangular, so that the covariance computation boils down to finding the variance of scalar uniform random variables.

Note that $B_k = 0$ implies $J_k = 0$, and together with $G_{\text{cl}}G_{\text{cl}}^T$, the above moments allow the computation of the linear optimal filter coefficients, K_k and L_k, according to (7) and (8), respectively.

C. Discussion

It is easy to see that, in the present case, $\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}$ is a block diagonal matrix such that the blocks along its main diagonal are

$$\begin{aligned}D = &\frac{1}{N}H_{\text{nom}}A\Lambda_kA^\top H_{\text{nom}}^\top + \frac{1}{N}H_{\text{nom}}\Sigma_{k+1}H_{\text{nom}}^\top \\ &+ \frac{1}{N}G_{\text{nom}}G_{\text{nom}}^\top + \frac{N-1}{N}G_{\text{cl}}G_{\text{cl}}^\top.\end{aligned} \tag{50}$$

In addition,

$$\begin{aligned}\Gamma_{x_{k+1}\tilde{y}_{k+1}} &= (\Sigma_{k+1} - A\Lambda_kA^\top)\mathbb{E}\left[H_{k+1}^\top\right] \\ &= \frac{1}{N}(\Sigma_{k+1} - A\Lambda_kA^\top)\Big(H_{\text{nom}}^\top \cdots H_{\text{nom}}^\top\Big)^\top.\end{aligned} \tag{51}$$

Therefore, the filter gain, K_k, in (7) becomes

$$\begin{aligned}K_k &= \Gamma_{x_{k+1}\tilde{y}_{k+1}}\Gamma_{\tilde{y}_{k+1}\tilde{y}_{k+1}}^{-1} \\ &= \frac{1}{N}\mathbf{1}_N^\top \otimes \Big((\Sigma_{k+1} - A\Lambda_kA^\top)H_{\text{nom}}^\top D^{-1}\Big).\end{aligned} \tag{52}$$

Bearing in mind that y_{k+1} is a concatenated vector of all the observations acquired at time $k+1$, the product K_ky_{k+1} in (2) is nothing but the average of the measurements constituting y_{k+1}, pre-multiplied by $(\Sigma_{k+1} - A\Lambda_kA^\top)H_{\text{nom}}^\top D^{-1}$. Consequently, the linear optimal estimator for tracking a target in clutter is a KF-like algorithm, operating on the average of all detections in the validation window. In this respect, its mode of operation resembles that of other classical methods. For example, the probabilistic data association (PDA) [19] method applies a KF on a weighted average of all measurements in the window. The weights are inversely proportional to the exponent of the Mahalanobis distance between each of the detections and the predicted measurement, corresponding to the covariance of the innovation. Similarly, the nearest neighbor (NN) technique [20], [21] applies a KF on the measurement nearest to the predicted estimate. This operation can be thought of as a weighted average in which one detection is assigned a weight of 1, while the rest are assigned 0 weights.

D. Numerical Study

In this section we demonstrate the performance of the linear optimal filter for tracking a target in clutter by comparing it with that of the NN and PDA methods. We consider a one-dimensional tracking scenario, in which a target is represented via a two dimensional state comprising position and velocity information, $x_k = (p_k,\, v_k)^\top$. Starting at $x_0 = (0,\, 0)^\top$ with $P_0 = \left(\begin{smallmatrix}0 & 0\\ 0 & 0\end{smallmatrix}\right)$, the target is simulated for 1000 time units by running the dynamical equation (39) with

$$A = \begin{pmatrix} 0.95 & 0.2 \\ 0 & 0.95 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{53}$$

The true measurement is generated by computing (40) with

$$H_{\text{nom}} = (1 \quad 0), \quad G_{\text{nom}} = 0.32. \tag{54}$$

For a (one-dimensional) window of length d, the clutter variance of (42) is $G_{\text{cl}}G_{\text{cl}}^\top = d^2/12$.

We compare the performance of our filter with that of the NN and PDA filters. All the considered algorithms are designed to reduce the mean-squared estimation error in either a heuristic or an analytical manner. However, when dealing with tracking in clutter, using the MSE as the only performance measure may result in misleading conclusions. This is due to the fact that once the estimated state draws away from the true measurement, clutter measurements are likely to be treated as the true ones, eventually resulting in target loss. When this happens, the algorithms' MSE becomes meaninglessly large. Therefore, the MSE should be treated as a meaningful performance measure only as long as the target is not lost. We use two measures

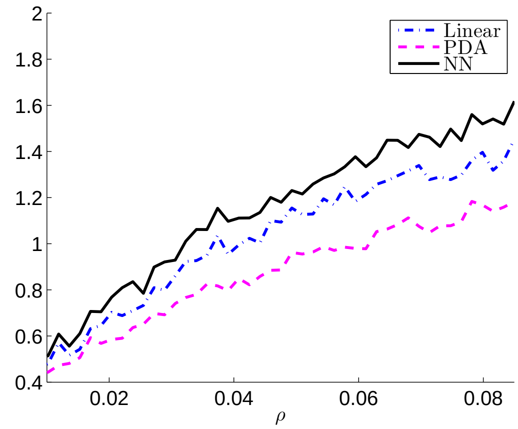

Fig. 2: Position RMSE vs. clutter density.

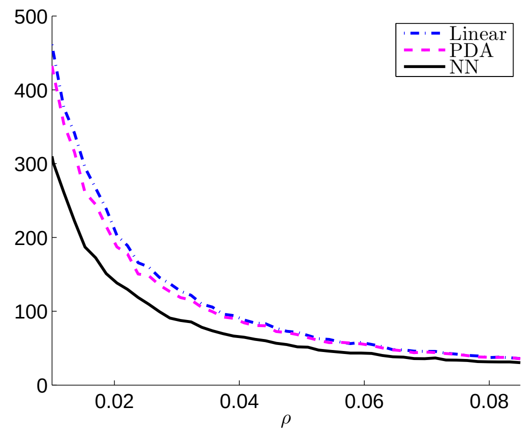

Fig. 3: Track loss times vs. clutter density.

of performance to evaluate the performance of the algorithms. The first is the time until the target is lost, defined as the third time the distance between the predicted position and the true state has deviated by more than five standard deviations of the (true) measurement noise. The second measure is the average squared error calculated over the time interval until the first of the three algorithms has lost track. This makes the comparison fair in the sense that none of the algorithms incorporates meaninglessly large errors corresponding to a lost target. A good algorithm is expected to have long track loss times, while maintaining low average squared errors.

We test the algorithms versus an increasing clutter density. To this end we define ρ to be the average number of clutter measurements falling in an interval of one standard deviation of the (true) measurement noise. Averaged over 1000 independent Monte Carlo runs, the average squared position errors and track loss times, versus ρ, are plotted in Figures 2, and 3, respectively.

It is readily seen that the linear optimal filter attains the longest track loss time while keeping the estimation errors at a reasonable compromise between the nonlinear PDA and NN algorithms.

V. Conclusion

We proposed a general formulation of dynamical systems with independently switching coefficients, where the dynamics and measurement equations are allowed to depend on previous estimates of the state. These additional terms may be used to represent closed-loop control input, and measurement validation window, respectively. We derived a linear optimal recursive algorithm for this setting, and illustrated the proposed approach in the context of target tracking in clutter. In this situation, the new filter demonstrates competitive performance, when compared with several classical nonlinear methods.

Acknowledgment

This research was supported by the Israeli Science Foundation under grant No. 1608/11 and the Technion's L. Kraus Research Fund.

References

[1] E.-K. Boukas and Z.-K. Liu, *Deterministic and Stochastic Time-Delay Systems*. Boston: Birkhäuser, 2002.
[2] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 457–462, July 1969.
[3] M. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Autom. Control*, vol. AC-24, no. 6, pp. 944–948, December 1979.
[4] R. Jackson and D. Murthy, "Optimal linear estimation with uncertain observations (Corresp.)," *IEEE Trans. Inf. Theory*, vol. 22, no. 3, pp. 376–378, 1976.
[5] D. Sigalov and Y. Oshman, "A new formulation of fault-tolerant estimation problems and some solutions," in *Proc. 26th Convention of Electrical and Electronis Engineers in Israel.*, 2011.
[6] N. Nahi and E. Knobbe, "Optimal linear recursive estimation with uncertain system parameters," *IEEE Trans. Autom. Control*, vol. 21, no. 2, pp. 263–266, 1976.
[7] G. Ackerson and K. Fu, "On state estimation in switching environments," *IEEE Trans. Autom. Control*, vol. 15, no. 1, pp. 10–17, 1970.
[8] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. 33, no. 8, pp. 780–783, 1988.
[9] D. Sigalov and Y. Oshman, "State estimation in hybrid systems with a bounded number of mode transitions," in *Proc. Fusion 2010, 13th International Conference on Information Fusion*, 2010.
[10] W. De Koning, "Optimal estimation of linear discrete-time systems with stochastic parameters," *Automatica*, vol. 20, no. 1, pp. 113–115, 1984.
[11] O. Costa, "Linear minimum mean square error estimation for discrete-time Markovian jump linear systems," *IEEE Trans. Autom. Control*, vol. 39, no. 8, pp. 1685–1689, 1994.
[12] T. Michaeli, D. Sigalov, and Y. Eldar, "Partially linear estimation with application to sparse signal recovery from measurement pairs," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2125–2137, 2012.
[13] T. Michaeli and Y. C. Eldar, "Hidden relationships: Bayesian estimation with partial knowledge," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 1933–1978, May 2011.
[14] D. Magill, "Optimal adaptive estimation of sampled stochastic processes," *IEEE Trans. Autom. Control*, vol. 10, no. 4, pp. 434–439, 1965.
[15] D. Sigalov and Y. Oshman, "Linear optimal estimation problems in systems with actuator faults," in *Proc. Fusion 2011, 14th International Conference on Information Fusion*, 2011.
[16] A. Jaffer and S. Gupta, "On Estimation of Discrete Processes Under Multiplicative and Additive Noise Conditions," *Information Sciences*, vol. 3, no. 3, pp. 267–276, 1971.
[17] J. Mendel, *Lessons in digital estimation theory*. Prentice-Hall, Inc., 1986.
[18] D. Sigalov, T. Michaeli, and Y. Oshman, "Linear optimal state estimation in systems with independent mode transitions," in *Proc. CDC 2011, 50th Conf. on Decision and Control*. IEEE, 2011.
[19] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.
[20] Y. Bar-Shalom and X. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.
[21] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Artech House, 1999.